REMARKS

Claims 1, 3-25, 27-30, 32-42, 44-48, 50-60, and 62-64 are currently pending in the subject application and are presently under consideration. Claims 1, 3-19, 21-25, 27-30, 32-38, 41-42, 44-48, 50-51, 53-57, 59-60, and 62-64 have been amended, and claims 2, 26, 31, 43, 49 and 61 have been cancelled without prejudice or disclaimer. A version of all pending claims is located at pages 2-22 of this Reply. Favorable reconsideration of the subject patent application is respectfully requested in view of the amendments and comments herein.

I. Rejection of Claims 1-64 Under 35 U.S.C. §101

Claims 1-64 stand rejected under 35 U.S.C. §101 as allegedly being directed towards non-statutory subject matter. Withdrawal of this rejection is requested for at least the following reasons. Independent claims 1, 19, 30, 42, 44, 53-54, and 62-64 have been amended herein to more clearly emphasize features of applicants' claimed subject matter. Accordingly, withdrawal of this rejection is requested.

II. Rejection of Claims 1-64 Under 35 U.S.C. §112

Claims 1-64 stand rejected under 35 U.S.C. §112, first paragraph, allegedly because current case law and the MPEP require such a rejection for claims that stand rejected under 35 U.S.C. §101. This rejection should be withdrawn for at least the following reasons. As stated above, independent claims 1, 19, 30, 42, 44, 53-54, and 62-64 have been amended to further emphasize aspects of the claimed subject matter. Accordingly, this rejection should be withdrawn.

III. Rejection of Claims 1, 19, 30, 42, and 64 Under 35 U.S.C. §102(b)

Claims 1, 19, 30, 42, and 64 stand rejected under 35 U.S.C. §102(b) as being anticipated by Guha *et al.* (US 5,140,530). Withdrawal of this rejection is requested for at least the following reasons. Guha *et al.* does not disclose or suggest all limitations set forth in the subject claims.

> A single prior art reference anticipates a patent claim only if it ***expressly or inherently describes each and every limitation set forth in the patent claim***. *Trintec Industries, Inc. v. Top-U.S.A. Corp.,* 295 F.3d 1292, 63 USPQ2d 1597 (Fed. Cir. 2002); *See Verdegaal Bros. v. Union Oil Co. of California,* 814 F.2d 628, 631, 2 USPQ2d 1051, 1053 (Fed. Cir. 1987). The ***identical invention must be shown in as complete detail as is contained in the ... claim***. *Richardson v. Suzuki Motor Co.,* 868 F.2d 1226, 9 USPQ2d 1913, 1920 (Fed. Cir. 1989) (emphasis added).

To date, where one has had access to massive amounts of data, the cost of building statistical models to characterize such data has been found to be significant if not insurmountable. The accuracy of the model and the cost of building the model are both competing interests associated with building a statistical model. Thus, while use of extremely large data sets may provide very accurate models in comparison to smaller sets of the data, analysis of data tends to become increasingly inefficient and expensive with larger data sets. Consequently, due to the computational complexities associated with analyzing large data sets, a common practice is to build models on the basis of a sample of the data. Nevertheless, the choice of sample size can be problematic.

Various methodologies have been proposed that employ progressive samples to analyze data in order to find an adequate sample size for which a model of reasonable quality can be constructed. A learning curve method (also known as progressive sampling) is one approach to evaluate the relationship between the accuracy of a model and the cost of learning the model. The basic idea being to iteratively apply a learning method to larger and larger subsets of data until the increasing costs of learning from larger subsets outweigh the increasing benefit of accuracy.

Applicants' claimed subject matter relates to a system and method to facilitate building a model to characterize data based on a subset of the data having an appropriate size. The claimed subject matter constructs a crude model for an initial subset of data using a first parameter estimation algorithm. The model may be evaluated, for example, by applying the model relative to a holdout data set of the data. Where the model is unacceptable, additional data can be added to the data subset and the first parameter estimation algorithm is repeated for the aggregate data subset. An appropriate subset of

the data exists when the first parameter estimation algorithm produces an acceptable model. The appropriate subset of the data subsequently can be employed by a different parameter estimation algorithm to build a statistical model that more accurately characterizes the data in its entirety. The subject matter as claimed in one aspect provides a relatively fast determination of an adequate size for the training data in situations where parameters will be estimated by employing a known parameter estimation technique (*e.g.*, an Expectation Maximization (EM) algorithm, and the like). To this end, independent claim 1 recites: ***a data scheduler that, based on a data policy, adaptively controls the size of subsets for which the first algorithm is applied***, independent claims 19, 30, 42, 54, and 62-64 recite: ***utilizes a stopping criterion that is functionally related to an expected incremental benefit and an expected incremental cost associated with increasing the size of the subset of data***, and independent claims 44 and 53 recite: ***controlling parameter initialization employed in each estimation of model parameters repeatedly until an acceptable size for the determined data subset is achieved***. Guha *et al.* does not disclose or suggest these aspects of the claimed subject matter.

Guha *et al.* relates to genetic learning techniques to evolve neural network architectures for applications where a general representation of neural network architecture is linked with a genetic learning strategy creating an environment for the construction of custom neural networks. In particular, the cited document involves the use of genetic algorithm methods to design new neural networks. The genetic algorithm (GA) is a robust function optimization method used in preference to gradient descent techniques for problems fraught with local minima, discontinuity, noise, or large numbers of dimensions. Guha *et al.* however does not adaptively control the size of the subset for which the first algorithm is applied, neither does the cited document employ a stopping criterion functionally related to an expected incremental benefit and an expected incremental cost associated with increasing the size of the subset of data, and further Guha *et al.* does not control parameter initialization at each estimation of model parameters until an acceptable size of the determined data subset is achieved. As can best be understood from the disclosure provided by Guha *et al.*, the cited document employs a learning algorithm that iteratively for each training-input and desired-output pair in a training set, applies the training-input to the input of a neural network, calculates the

output from the network, determines whether or not the calculated output is equal to the desired-output, if the calculated output is not equal to the desired-output network weights associated with the neural network are modified. Each iteration of the learning algorithm is referred to as an epoch, and two termination conditions can be used: a lower bound on the error over an epoch and a limit on the number of epoch. However, as stated *supra*. the cited document is silent with respect to the salient features recited in the subject independent claims. Accordingly, in view of at least the foregoing withdrawal of this rejection is requested.

CONCLUSION

The present application is believed to be in condition for allowance in view of the above amendments and comments. A prompt action to such end is earnestly solicited.

In the event any fees are due in connection with this document, the Commissioner is authorized to charge those fees to Deposit Account No. 50-1063 [MSFTP184US].

Should the Examiner believe a telephone interview would be helpful to expedite favorable prosecution, the Examiner is invited to contact applicants' undersigned representative at the telephone number below.

Respectfully submitted,

AMIN, TUROCY & CALVIN, LLP

/Himanshu S. Amin/
Himanshu S. Amin
Reg. No. 40,894

AMIN, TUROCY & CALVIN, LLP
24TH Floor, National City Center
1900 E. 9TH Street
Cleveland, Ohio 44114
Telephone (216) 696-8730
Facsimile (216) 696-8731